



Nucleotide

My NCBI
[Sign In] [Register]

All Databases

PubMed

Nucleotide

Protein

Genome

Structure

OMIM

PMC

Journals

Books

Search for

Go

Clear

[Limits](#) [Preview/Index](#) [History](#) [Clipboard](#) [Details](#)[Format:](#) [GenBank](#) [FASTA](#) [Graphics](#) [More Formats▼](#) [Download▼](#) [Save▼](#) [Links▼](#)

GenBank AY274119.3

SARS coronavirus TOR2, complete genome[Change Region Shown](#)[Customize View](#)[Comment](#) [Features](#) [Sequence](#)[Analyze This Sequence](#)

LOCUS AY274119 29751 bp RNA linear VRL

24-MAR-2004

DEFINITION SARS coronavirus TOR2, complete genome.

ACCESSION AY274119

VERSION AY274119.3 GI:30248028

KEYWORDS

SOURCE SARS coronavirus Tor2

ORGANISM SARS coronavirus Tor2

Viruses; ssRNA positive-strand viruses, no DNA stage;

Nidovirales;

Coronaviridae; Coronavirinae; Betacoronavirus.

REFERENCE 1 (bases 1 to 29751)

AUTHORS He,R., Dobie,F., Ballantine,M., Leeson,A., Li,Y.,

Bastien,N.,

Cutts,T., Andonov,A., Cao,J., Booth,T.F., Plummer,F.A.,

Tyler,S.,

Baker,L. and Li,X.

CONSRTM BCCA Genome Sciences Centre, British Columbia Centre
for Disease

Control and National Microbiology Laboratory Canada

TITLE Analysis of multimerization of the SARS coronavirus

nucleocapsid

protein

JOURNAL Biochem. Biophys. Res. Commun. 316 (2), 476-483 (2004)

PUBMED 15020242

REFERENCE 2 (bases 1 to 29751)

CONSRTM BCCA Genome Sciences Centre, British Columbia Centre

for Disease

Control and National Microbiology Laboratory Canada

TITLE Direct Submission

JOURNAL Submitted (13-APR-2003) Genome Sciences Centre, British

Columbia

Cancer Research Centre, 600 West 10th Avenue,

Vancouver, BC V5Z

4E6, Canada

REFERENCE 3 (bases 1 to 29751)

CONSRTM BCCA Genome Sciences Centre, British Columbia Centre

for Disease

Control and National Microbiology Laboratory Canada

TITLE Direct Submission

JOURNAL Submitted (23-APR-2003) Genome Sciences Centre, British

Columbia

Cancer Research Centre, 600 West 10th Avenue,

Vancouver, BC V5Z

4E6, Canada

REMARK Sequence update by submitter

REFERENCE 4 (bases 1 to 29751)

CONSRTM BCCA Genome Sciences Centre, British Columbia Centre

for Disease

Control and National Microbiology Laboratory Canada

[Run BLAST](#)[Pick Primers](#)**SARS Coronavirus Resource**Compiled data and information
relevant to SARS coronavirus.**Recent activity**[Turn Off](#) [Clear](#)Laboratory safety aspects
of SARS at Biosafety
Level 2.SARS coronavirus TOR2,
complete genomeSARS coronavirus TOR2,
complete genomeSARS coronavirus TOR2,
complete genome

AY274119 (1)

[» See more...](#)**All links from this record**[Related sequences](#)[Full text in PMC](#)[Genome](#)[Other nucleotides for species](#)[Probe](#)[Protein](#)[PubMed](#)

TITLE Direct Submission
JOURNAL Submitted (30-APR-2003) Genome Sciences Centre, British Columbia
Cancer Research Centre, 600 West 10th Avenue,
Vancouver, BC V5Z
4E6, Canada

PubMed (weighted)

RefSeq nucleotide for species

Taxonomy

REMARK Sequence update by submitter
COMMENT On Apr 30, 2003 this sequence version replaced
gi:30088476.

FEATURES
 source
 Location/Qualifiers
 1..29751
 /organism="SARS coronavirus Tor2"
 /mol_type="genomic RNA"
 /isolate="Tor2"
 /isolation_source="patient #2 with severe acute
 respiratory syndrome (SARS)"
 /db_xref="taxon:227984"
 /country="Canada: Toronto"
 join(265..13392,13392..21485)
 /ribosomal_slippage
 /codon_start=1
 /product="replicase 1AB"
 /protein_id="AAP41036.1"
 /db_xref="GI:30795144"
 /translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGD

SVEEALSEARE

HLKNGTCGLVELEKGVLPQLEQPYVFIKRS DALSTNHGHKVV ELVAEMDGIQYGRSGI
TLGVLVPHVGETPIAYRNVLLRKNGK GAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
WNTKHGSGALREL TRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAKKFDTFKGECPKFVFP
LNSKVVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMFPCACQDPEIGFEHSVADYHNHNSN
IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
LEILSRERVNINIVGDFHLNEEVAILASFSASTSAFIDTIKSLDYKSFKTIVESCNG
YKVTKGKPKVGAWNIGQQRSVLTPLCGFPSQAAGVIRSI FARTLDAANH SIPDLQRAA
VTILDGISEQSLRLVDAMVYTSDLLTNSVIMAYVTGGLVQQT SQWLSNLLGTTVEKL
RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCVKCFIDVV
NKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKGLYRQCIRGKEQLQLLMPLKAPKEVT
FLEGDSHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGLMLEIKDKEQ
YCALSPGLLATNNVFRLLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE
KCSVYTVESGTEVTEFACVVAEAVVKTLPVSDLLTNMGIDLD EWSVATFYLFDDAGE
ENFSSRMYSFYPDEEEEDDAECBEEEEIDETCEHEYGTEDDYQGLPLEFGASAETVR
VEEEEEEDWLD DTEQSEIEPEPEPTPEEPVNQFTGYLKLTDNVAIKCVDIVKEAQA
NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
AKKCLHVVGPNNLAGEDIQLLKAAYENFNSQDILLAPILSAGIFGAKPLQSLQVCVQT
VRTQVYIAVNDKALYEQVVM DYLDNLKPRVEAFKQEEFPNTEDSKTEEKSVVQKFPVDV
KPKIKACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEKDAPYM
VGDVITSGDITCVVIPS KAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTLERAKT
ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEETRKLMPICMDVRAIMATI

QRKYKGIKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
EAARCMRSLKAPAVVSVSSPDAVTTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSG
QRTELGVFEFLKRGRKIVYHTLESFVEFHLDEVLSDKLKSLLSREVKTIKVFTTVD
NTNLHTQLVDMSMTYGGQPGPTYLDGADVTKIKPHVNHEGKTFVLPSSDTRLSEAFE
YYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQQLEVKN
APALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHTLLQHANLESKRVLN
VVCKHCGQKTTTLTGVEAVMYMGTLSDNLKTGVSIPCVCGRATQYLVQQESSFVMM
SAPPAEYKLOQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD
VFYKETSytttikpvsykldgvtYTEIEPKLDGYKKDNAYYTEQPIDLVPTQPLPNA
SFDNFKLTCSNTRFADDLNQMTGFTKPAARELSVTFFPDNLGDVVAIDYRHSASPKK
GAKLLHKPIVWHINQATTKTTFKPNWCLRLWSTKPVDTSSNFEVLAVEDTQGMNDL
ACESQQTSEEVENPTIQKEVIECDVKTTEVVGNVILKPSDEGVKVTQELGHEDLMA
AYVENTSITIKKPNELSLALGLKTIATHGIAAINSVPWSKILAYVKPFLGQAAITTSN
CAKRLAQRVFNMYPYVFTLLFQLCTFTKSTNSRIRASLPTTIKNSVKSVAKLCLDA
GINYVKSPEKFLFTIAMWLLLLSICLSLICVTAAGVLLSNFGAPSYCNGVRELYL
NSSNVTTMDFCEGSFPCSICLSGLDSDSYPALETIQVTISSYKLDLTILGLAAEWL
AYMLFTKFFYLLGLSAIMQVFPGYFASHFISNSWLMWFIISIVQMAPVSAMVRMYIFF
ASFYYIWKSIVHTMDGCTSSCTMMCYKRNRRATRVECTIVNGMKRSFYVYANGGRGFC
KTHNWNCLNCDTCTGSTFISDEVARDLSLQFKRPINPTDQSSYIVDSVAVKNGALHL
YFDKAGQKTYERHPLSHFVNLDNLRANNTKGSPLINVIVFDGKSKCEDASASKSASVYY
SQLMCQFILLLDQALVSDVDGSTEVSVKMFDAYVDTFSATFSVPMEKLLKALVATAHSE
LAKGVALDGVLSFVSAARQGVVDTVDTKDVIECLKLSHSDLEVTGDSCNNFMLTY
NKVENMTPRDLGACIDCNARHINAQVAKSHNVSLIWNVKDYMSLSEQLRKQIRSAKK
NNIPFRLTCATTRQVVNVITTKISLKGKIVSTCFKMLKATLLCVLAALVCYIVMPV
HTLSIHDGYTNEIIGYKAIQDGVTRDIISTDDCFANKHAGFDWFSQRGGSYKNDKSC
PVVAAIITREIGFIVPGLPGTVLRAINGDFLHFLPRVFSAVGNICYTFSKLEYSDF
TSACVLAAECTIFKDAMGKVPYCYDTNLLSGSISYSELRPDTRYVLMGDSIIQFPNT
YLEGSRVVVTTFDAEYCRHGTCESEVIGICLSTSGRWVLNNEHYRALSGVFCGV DAMN
LIANIFTPLVQPVGALDVSAVAGGIIAILVTCAAYYFMKFRRVFGEYNHVVAANAL
LFLMSFTILCLVPAYSFLPGVYSVFYLYLTFYFTNDVSFLAHLQWFAMFSPIVPFWIT
AIYVFCISLKHCHWFFNNYLKRVMFNGVTFSTFEAAALCTFLNKEMYLKLRSETLL
PLTQYNYRLALYNKYKYFGALDTSYREACCHLAKALNDFSNGADVLYQPPQTSI
TSAVLQSGFRKMAFSPGKVEGCMVQVTCGTTTLNGLWLDLTVYCPRHVICTAEDMLNP
NYEDLLIRKSNHSFLVQAGNVQLRVIGHSMQNCLLRLKVDTSNPKTKYKVFVRIQPGQ
TFSVLACYNGSPSGVYQCAMPNHTIKGSFLNGSCGSGVFNIDYDCVSFCYMHMELP
TGVHAGTDLEGKFYGPVDRQTAQAAGTDTTITLNLAWLYAAVINGDRWFLNRF TTT

[Write to the Help Desk](#)

[NCBI](#) | [NLM](#) | [NIH](#)

[Department of Health & Human Services](#)
[Privacy Statement](#) | [Freedom of Information](#)
[Act](#) | [Disclaimer](#)

LNDNFLVAMKYNIEPLTQDHVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGS
TILEDEFTFPDVRQCSGVTGQKFKKIVKGTTHHMLLTFLTSLILVQSTQWSLFFF
VYENAFLPFTLGIMAIAACAMLLVKHKHAFCLCLLPSLATVAYFNMVYPASWVMRI
MTWLELADTSLSGYRLKDCVMYASALVLLILMTARTVYDDAARRVWTLMNVITLVYKV
YYGNALDQAISMWALVISVTSNYSYGVVTTIMFLARAIIVFCVEYYPLLFITGNTLQCI
MLVYCFGLGYCCCCYFGLFCLLNRYFRLTLGVYDYLSTQEFYMNSSQGLLPKSSIDA
FKLNIKLLGIGGKPKIKVATVQSKMSDVKCTSVVLLSVLQQLRVESSSKLWAQCVQLH
NDILLAKDTTEAFEKMSVLLSVLLSMQGAVIDINRLCEEMLDNRATLQAIASEFSSSLPS
YAYATAQEAYEQAVANGDSEVVLKLLKSLNVAKSEFDRDAAMQRKLEKMAQAMTQ
MYKQARSEDKRAKVTSAMQTMFLTLRKLDNDALNNIINNARDGCVPLNIIPLTTAAK
LMVVVPDYGTYNKTCGNTFTYASALWEIQQVVDADSKIVQLSEINMDNSPNLAWPLI
VTALRANSVKLQNNELSPVALRQMSCAAGTTQACTDDNALAYNNKSGGRFVLALL
SDHQDLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGPVKYLYFIKGLNNLNRGMVL
GSLAATVRLQAGNATEVPANSTVLSFCAFAVDPAKAYKDYLASGGQPITNCVKMLCTH
TGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNPKGFCDLKGKYVQIPTTCAND
PVGFTLRNTVCTVCGMWKGYGCSCDQLREPLMQSADASTFFKRVCGVSAARLTPCGTG
TSTDVVYRAFDIYNEKVAGFAKFLKTNCCRFQEKDEEGNLLDSYFVVKRHTMSNYQHE
ETIYNLVKDCPAVAVHDFPKFRVDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCDT
LKEILVTYNCCDDDYFNKKDWYDFVENPDILRVYANLGERVRQSLLKTVQFCDAMRDA
GIVGVLTLDNQDLNGNWYDFGDFVQVAPGCGVPIVDSYSLMLPILTLTRALAAESHM
DADLAKPLIKWDLKYDFTEERLCLFDRYFKYWDQTYHPNCINCLDDRCILHCANFNV
LFSTVFPPTSFGPLVRKIFVDGVPFVSTGYHFRELGVVHNQDVNLHSSRLSFKELLV
YAADPAMHAASGNLLLDKRTTCFSAALTNNVAFQTVKPGNFNKDFYDFAVSKGFFKE
GSSVELKHFFFAQDGNAAISDYDYRYNLFMTCDIRQLLFVVEVVDKYFDCYDGGCIN
ANQVIVNNLDSAGFPFNKWKARLYYDSMSYEDQDALFAYTKRNVIPITITQMNLYA
ISAKNRARTVAGVSICSTMNRQFHQKLLKSAATRGATVVIGTSKFYGGWHNMLKTV
YSDVETPHLMGWDPKCDRAMPNMLRIMASLVLARKHNTCCNLSHRFYRLANCAQVL
SEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYVR
NLQHRLYECLYRNRDVEHFVDEFYAYLRKHFSMMILSDDAVVCYNSNYAAQGLVASI
KNFKAVLYYQNNVMSEAKCWTETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRI
LGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADVFLYLQYIRKLHDEL
TGHMLDMYSVMLTNDNTSRYWEFEFYAMYTPHTVLQAVGACVLCNSQTSRLCGACIR
RPFLCCKCCYDHVISTSHKLVLVSNPYVCNAPGCDVTDVTQLYLGGMSYCKSHKPP
SFPLCANGQVFGLYKNTCVGSNDVTDFNAIATCDWTNAGDYILANTCTERLKLFAAET
LKATEETFKLSYGIATVREVLSDRELHLSWEVGKRPPLNRNYVFTGYRVTKNSKVQI
GEYTFEKGDYGDVAVYRGITTYKLVNGDYFVLTSHTVMPLSAPTLVPQEHYVRITGLY

PTLNISDEFSSNVANYQKVGMMQKYSTLQGPPGTGKSHFAIGLALYYPSARIVYTACSH
AAVDALCEKALKYLPIDKCSRIIPARARVECFDKFKVNSTLEQYVFCTVNALPETTAD
IVVFDEISMATNYDLSVNVNARLRKHYVYIGDPAQLPAPRTLLTKGTLEPEYFNSVCR
LMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLKAHKDKSAQCFKMFYKGVITHDVS
SAINRPQIGVVRFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVDSSQGSEYDY
VIFTQTTETAHSCNVNRFNVAITRAKIGILCIMSDDRDLYDKLQFTSLEIPRNVATLQ
AENVTLGLFKDCSKIITGLHPTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLISMM
GFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEGCHATRDAVGTNLPLQLGFSTGVN
LVAVPTGYVDTENNTFTRVNAKPPPGDQFKHLIPLMYKGLPWNVVRKIVQMLSDTL
KGLSDRVVFLVLAHGFELTSMKYFVKIGPERTCCLCDKRATCFSTSSDTYACWNHSVG
FDYVYNPFMIDVQQWGFCTGNLQSNHDQHCQVHGNAHVASCDAIMTRCLAVHECFVKRV
DWSVEYPIIGDELRVNSACRKVQHMVKSALLADKFPVLHDIGNPKAIKCPQAEVEW
KFYDAQPCSDKAYKIEELFYSYATHHDKFTDGVCLFWNCNVDRYPANAIVCRFDTRVL
SNLNLPGCDGGSLYVNKHAFHTPAFDKSAFTNLKQLPFFYYSDSPCESHGKQVVSIDID
YVPLKSATCITRCNLGGAVCRHHANEYRQYLDAYNMMISAGFSLWIYKQFDTYNLWNT
FTRLQSLENVAYNVNKGHFDGHAGEAPVSIINNAVYTKVDGIDVEIFENKTTLPVNV
AFELWAKRNKIPVPEIKILNGLVDIAANTVIWDYKREAPAHVSTIGVCTMTDIAKKP
TESACSSLTVLFDGRVEGQVDLFRNARNGVLITEGSVKGLTPSKGPAQASVNGVTLIG
ESVKTQFNYFKKVDGIIQQLPETYFTQSRDLEDFKPRSQMETDFLELAMDEFIQRYKL
EGYAFEHIVYGDPSHGQLGGLHLMIGLAKRSQDSPLKLEDFIPMDSTVKNYFITDAQT
GSSKCVCSVIDLLLDFFVEIISQDLSVISKVVKTIDYAEISFMLWCKDGHVETFP
KLQASQAWQPGVAMPNLYKMQRMLLEKCDLQNYGENAVIPKGINMNAKYTQLCQYLN
TLTLAVPYNMRIHFGAGSDKGVA PGTA VL RQWLPTGTL LVDS DLNDFVSDADSTLIG
DCATVHTANKWDLIISDMYDPRTRKHVTKENDSKEGFFTYLCGFIKQKLALGGSIAVKI
TEHSWNADLYKLMGHFSWWTAFVTNVNASSSEAFILGANYLGKPKQIDGYTMHANYI
FWRNTNPIQLSSYSFLDMSKFPKLKRGTA VMSLKENQINDMIYSLLEKGRLI IRENNR
VVSSDILVNN"
21492..25259
/codon_start=1
/product="spike glycoprotein"
/protein_id="AAE31037.1"
/db_xref="GI:30795145"
/translation="MFIFLLFLTLTSGSDLRCTTFDDVQAPNYTQH
TSSMRGVYYPD
EIFRSDTLYLTDQLFLPFYSNVTGFHTINHTFGNPFVIPFKDGIYFAATEKSNVVRGWV
FGSTMNKSQSVIIINNSTNVVIRACNFELCDNPFPAVSKPMGTQHTMIFDNAFNCT
FEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFYVYKGYQPIDVVRDLPSGFNTLKP
IFKLPLGINITNFRAILTAFSPAQDIWGTSAAYFVGYLKPTTFMLKYDENGITIDAV
DCSQNPALAEKCSVKSFEIDKGIYQTSNFRVVPVSGDVVRFPNITNLCPFGEVFNATKF
PSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNLCFSNVYADS FVVKGD

DVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRP
FERDISNVPFSPDGKPCPPALNCYWPLNDYGFYTTTIGIGYQPYRVVLSFELLNAPA
TVCGPKLSTDLIKNQCVNPNFNGLTGTGVLPSSKRFQPFQFGRDVSDFDTSVRDPK
TSEILDISPFAFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYST
GNNVFQQTQAGCLIGAHEVDTSYECDIPIGAGICASYHTVSLRSTSQKSIVAYTMSLG
ADSSIAYSNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGS
FCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRS
FIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDMDIAAYT
AALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAIS
QIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRDKVEAE
VQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDPCGKGYH
LMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVVFVNGTSWFITQ
RNFFSPQIITDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDV
DLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFA
GLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKFDEDDSEPVLKGVKLHYT"
CDS 25268..26092
/codon_start=1
/product="Orf3"
/protein_id="AAP41039.1"
/db_xref="GI:30795146"
/translation="MDLFMRFFTLRSITAQPVKIDNASPASTVHATA
TIPLQASLPFG
WLIVIGVAFLAVFQSATKIIALNKRWQLALYKGFQFICNLLLLFVTIYSHLLLVAAGME
AQFLYLYALIIYFLQCINACRIIMRCWLCKWCKSKNPLLYDANYFVCWHTHNYDYCIPY
NSVTDITIVVTEGDGISTPKLKEDYQIGGYSEDRHSGVKDYVVVHGYFTEVYYQLESTQ
ITTDGTGIENATFFIFNKLVDPPNVQIHTIDGSSGVANPAMDPIYDEPTTTTSVPL"
CDS 25689..26153
/codon_start=1
/product="Orf4"
/protein_id="AAP41039.1"
/db_xref="GI:30795147"
/translation="MMPTTLFAGTHITMTTVYHITVSQIQLSLLKVT
AFQHQNKKTT
KLVVILRIGTQVLKTMSTLYMAISPKFTTSLSLHKLQLTLVLKMLHSSSLTSLLKTHRM
CKYTQSTALQELLIQQWIQFMMSRRRLACLCKHKKVSTNLCTHSFRKKQVR"
CDS 26117..26347
/codon_start=1
/product="small envelope E protein"
/protein_id="AAP41040.1"
/db_xref="GI:30795148"
/translation="MYSFVSEETGTLIVNSVLLFLAFVVFLLVTLAI
LTALRLCAYCC
NIVNVSLVKPTVYVYSRVKNLNSSEGVPDLLV"
CDS 26398..27063
/codon_start=1
/product="membrane glycoprotein M"
/protein_id="AAP41041.1"
/db_xref="GI:30795149"
/translation="MADNGTITVEELKQLLEQWNLVIGFLFLAWIML
LQFAYSNNRRF

LYIIKLVFLWLLWPVTLACFVLAAYVRINWVTGGIAIAMACIVGLMWLSYFVASFRLLF
ARTRSMWSFNPETNILLNVPLRGITVTRPLMESELVIGAVIIRGHLRMAGHSLGRCDI
KDLPEKITVATSRTLSYYKLGASQVRGTDGFAAYNRYRIGNYKLNTDHAGSNDNIAL
LVQ"
CDS 27074..27265
/codon_start=1
/product="Orf7"
/protein_id="AAP41042.1"
/db_xref="GI:30795150"
/translation="MFHLVDFQVTIAEILIIIMRTFRIAIWNLDVII
SSIVRQLFKPL
TKKNYSELDDDEEPMELDYP"
CDS 27273..27641
/codon_start=1
/product="Orf8"
/protein_id="AAP41043.1"
/db_xref="GI:30795151"
/translation="MKIILFLTLIVFTSCELYHYQECVRGTTVLLKE
PCPSGTYEGNS
PFHPLADNKFALTCTSTHFAFACADGTRHTYQLRARSVSPKLFIRQEEVQQELYSPLF
LIVAALVFLILCFTIKRKTE"
CDS 27638..27772
/codon_start=1
/product="Orf9"
/protein_id="AAP41044.1"
/db_xref="GI:30795152"
/translation="MNELTLIDFYLCFLAFLFLVLIMLIIFWFSLE
IQDLEEBCTKV"
CDS 27779..27898
/codon_start=1
/product="Orf10"
/protein_id="AAP41045.1"
/db_xref="GI:30795153"
/translation="MKLLIVLTCISLCSICTVVQRCASNKPHVLED
PCKVQH"
CDS 27864..28118
/codon_start=1
/product="Orf11"
/protein_id="AAP41046.1"
/db_xref="GI:30795154"
/translation="MCLKILVRYNTRGNTYSTAWLCAKGLVPFHRW
HTMVQTCTPNV
TINCQDPAGGALIARCWYLHEGHQTAAFRDVLVVLNKRNT"
CDS 28120..29388
/codon_start=1
/product="nucleocapsid protein"
/protein_id="AAP41047.1"
/db_xref="GI:30795155"
/translation="MSDNGPQSNQRSAPRITFGGPTDSTDNNQNGGR
NGARPKQRRPQ
GLPNNTASWFTALTQHGKEELRFPRGQVPINTNSGPDQIGIYYRRATRRVRGGDGKM
KELSPRWYFYLLGTGPEASLPYGANKEGIVWVATEGALNTPKDHIGTRNPNNAATVL
QLPQGTTLPLKGFYAEGSRGGSQASSRSSSRGNSRNSTPGSSRGNSPARMASGGGET
ALALLLLDRLNQLSEKVSQKGGQQGGQTVTKKSAAEASKKPRQKRTATKQYNVTQAFG
RRGPEQTQGNFGDQDLIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTY
HGAIKLDDKDPQFKDNVILLNKHIDAYKTFPPTEPKKDKKKKTDEAQPLPQRQKKQPT
VTLPLAADMDDFSRQLQNSMSGASADSTQA"
CDS 28130..28426
/codon_start=1
/product="Orf13"
/protein_id="AAP41048.1"
/db_xref="GI:30795156"
/translation="MDPNQTNVPPALHLVDPQIQLTITRMEDAMGQ

GQNSADPKVYP

IILRLGSQLSLSMARRNLDSEARAFQSTPIVVQMTKLATTEELPDEFVVTAK"

CDS

28583..28795

/codon_start=1

/product="Orf14"

/protein_id="AAP_1049_1"

/db_xref="GI:30795157"

/translation="MLFPCYNFLKEQHCQKASTQREAEAAVKPLLPAP

HHVVAVIQEIQ

LLAAVGEILLLEWLAIEVKLPSPRYCC"

ORIGIN

```
1 atattaggtt tttacctacc caggaaaagc caaccaacct cgatctcttg
tagatctggt
61 ctctaaacga actttaaaat ctgtgtagct gtcgctcggc tgcattgcta
gtgcacctac
121 gcagtataaa caataataaa ttttactgtc gttgacaaga aacgagtaac
tcgtccctct
181 tctgcagact gcttacgggt tcgtccgtgt tgcagtcgat catcagcata
cctagggttc
241 gtccgggtgt gaccgaaagg taagatggag agccttggtc ttggtgtcaa
cgagaaaaca
301 caggtccaac tcagtttgcc tgctcttcag gttagagacg tgctagtgcg
tggcttcggg
361 gactctgtgg aagaggccct atcggaggca cgtgaacacc tcaaaaatgg
cacttggtgt
421 ctagtagagc tggaaaaagg cgtactgccc cagcttgaa acgcctatgt
gttcattaaa
481 cgttctgatg ccttaagcac caatcacggc cacaaggctg ttgagctggt
tgcagaaatg
541 gacggcattc agtacggctg tagcgggtata aactggggag tactcgtgcc
acatgtgggc
601 gaaaccccaa ttgcataccg caatgttctt cttcgtaaga acggtataaa
gggagccggg
661 ggtcatagct atggcatcga tctaaagtct tatgacttag gtgacgagct
tggcactgat
721 cccattgaag attatgaaca aaactggaac actaagcatg gcagtgggtc
actccgtgaa
781 ctactcgtg agctcaatgg aggtgcagtc actcgtctatg tcgacaacaa
tttctgtggc
841 ccagatgggt accctcttga ttgcatcaaa gattttctcg cagcgcggg
caagtcaatg
901 tgcactcttt ccgaacaact tgattacatc gagtcgaaga gaggtgtcta
ctgctgccgt
961 gaccatgagc atgaaattgc ctgggttact gagcgtctg ataagagcta
cgagcaccag
1021 acacccttcg aaattaagag tgccaagaaa ttgacactt tcaaagggga
atgcccacaa
1081 tttgtgtttc ctcttaactc aaaagtcaaa gtcattcaac cacgtgttga
aaagaaaaag
1141 actgagggtt tcatggggcg tatacgtctt gtgtacctg ttgcatctcc
acaggagtgt
1201 aacaatatgc acttgtctac cttgatgaaa tgtaatcatt gcgatgaagt
ttcatggcag
1261 acgtgcgact ttctgaaagc cacttgtaga cattgtggca ctgaaaattt
agttattgaa
1321 ggacctacta catgtgggta cctacctact aatgctgtag tgaatatgcc
atgtcctgcc
1381 tgtcaagacc cagagattgg acctgagcat agtggttcag attatcacia
ccactcaaac
1441 attgaaactc gactccgcaa gggaggtagg actagatggt ttggaggctg
tgtgtttgcc
1501 tatgttggtt gctataataa gcgtgcctac tgggttcctc gtgctagtgc
tgatattggc
1561 tcaggccata ctggcattac tggtagaat gtggagacct tgaatgagga
tctccttgag
1621 atactgagtc gtgaacgtgt taacattaac attgttggcg attttcattt
gaatgaagag
1681 gttgccatca ttttggcacc tttctctgct tctacaagtg cctttattga
cactataaag
1741 agtcttgatt acaagtcttt caaaaccatt gttgagtcct gcggtaacta
taaagttacc
```


1801 aagggaagc ccgtaaaagg tgcttgaac attggacaac agagatcagt
tttaacacca
1861 ctgtgtggtt ttccctcaca ggctgctggg gttatcagat caatttttgc
gcgcacactt
1921 gatgcagcaa accactcaat tcttgatttg caaagagcag ctgtcaccat
acttgatggg
1981 atttctgaac agtcattacg tcttgctgac gccatgggtt atacttcaga
cctgctcacc
2041 aacagtgtca ttattatggc atatgtaact ggtgggtctg tacaacagac
ttctcagtgg
2101 ttgtctaata ttttgggcac tactgttgaa aaactcaggc ctatcttga
atggattgag
2161 gcgaaactta gtgcaggagt tgaatttctc aaggatgctt gggagattct
caaatttctc
2221 attacagggtg tttttgacat cgtcaagggt caaatacagg ttgcttcaga
taacatcaag
2281 gattgtgtaa aatgcttcat tgatgttgtt aacaaggcac tcgaaatgtg
cattgatcaa
2341 gtactatcgt ctggcgcaaa gttgcgatca ctcaacttag gtgaagtctt
catcgtctca
2401 agcaagggtg tttaccgtca gtgtatacgt ggcaaggagc agctgcaact
actcatgcct
2461 cttaaggcac caaaagaagt aacctttctt gaagggtgatt cacatgacac
agtacttacc
2521 tctgaggagg ttgttctcaa gaacggtgaa ctgaagcac tcgagacgcc
cgttgatagc
2581 ttcacaaatg gagctatcgt tggcacacca gtctgtgtaa atggcctcat
gctcttagag
2641 attaaggaca aagaacaata ctgcgcttgg tctcctggtt tactggctac
aaacaatgtc
2701 tttcgcttaa aaggggggtg accaattaaa ggtgtaacct ttggagaaga
tactgtttgg
2761 gaagttcaag gttacaagaa tgtgagaatc acatttgagc ttgatgaacg
tgttgacaaa
2821 gtgcttaatg aaaagtgtc tgtctacact gttgaatccg gtaccgaagt
tactgagttt
2881 gcatgtgttg tagcagaggc tgttgtaag actttacaac cagtttctga
tctccttacc
2941 aacatgggta ttgatcttga tgagtggagt gtagctacat tctacttatt
tgatgatgct
3001 ggtgaagaaa acttttcac acgtatgtat tgttcctttt accctccaga
tgaggaagaa
3061 gaggacgatg cagagtgtga ggaagaagaa attgatgaaa cctgtgaaca
tgagtacggg
3121 acagaggatg attatcaagg tctccctctg gaatttggtg cctcagctga
aacagtctga
3181 gttgaggaag aagaagagga agactggctg gatgatacta ctgagcaatc
agagattgag
3241 ccagaaccag aacctacacc tgaagaacca gttaatcagt ttactgggta
tttaaaactt
3301 actgacaatg ttgccattaa atgtgttgac atcgtaagg aggcacaaag
tgctaactct
3361 atgggtgattg taaatgtctg taacatacac ctgaaacatg gtgggtgggtg
agcaggtgca
3421 ctcaacaagg caaccaatgg tgccatgcaa aaggagagtg atgattacat
taagctaaat
3481 ggccctctta cagtaggagg gtcttgtttg ctttctggac ataactctgc
taagaagtgt
3541 ctgcatgttg ttggacctaa cctaaatgca ggtgaggaca tccagcttct
taaggcagca
3601 tatgaaaatt tcaattcaca ggacatctta cttgcaccat tgtgtcagc
aggcatattt
3661 ggtgctaaac cacttcagtc tttacaagtg tgcgtgcaga cggttcgtac
acaggtttat
3721 attgcagtca atgacaaagc tctttatgag cagggtgtca tggattatct
tgataacctg
3781 aagcctagag tggaagcacc taaacaagag gagccaccaa acacagaaga
ttccaaaact
3841 gaggagaaat ctgtcgtaca gaagcctgtc gatgtgaagc caaaaattaa
ggcctgcatt
3901 gatgagggtta ccacaacact ggaagaaact aagttttctta ccaataagtt
actcttggtt

3961 gctgatatca atggtaagct ttaccatgat tctcagaaca tgcttagagg
tgaagatatg
4021 tctttccttg agaaggatgc accttacatg gtaggtgatg ttatcactag
tggtgatatac
4081 acctgtgttg taataccctc caaaaaggct ggtggcacta ctgagatgct
ctcaagagct
4141 ttgaagaaag tgccagtga tgagtatata accacgtacc ctggacaagg
atgtgctggt
4201 tatacacttg aggaagctaa gactgctctt aagaaatgca aatctgcatt
ttatgtacta
4261 ccttcagaag cacctaattgc taaggaagag attctaggaa ctgtatcctg
gaatttgaga
4321 gaaatgcttg ctcatgctga agagacaaga aaattaatgc ctatatgcat
ggatgttaga
4381 gccataatgg caaccatcca acgtaagtat aaaggaatta aaattcaaga
gggcacggt
4441 gactatggtg tccgattctt cttttatact agtaaagagc ctgtagcttc
tattattacg
4501 aagctgaact ctctaaatga gccgcttgtc acaatgccaa ttggttatgt
gacacatggt
4561 tttaatcttg aagaggctgc gcgctgtatg cgttctctta aagctcctgc
cgtagtgtca
4621 gtatcatcac cagatgctgt tactacatat aatggatacc tcacttcgtc
atcaaaagaca
4681 tctgaggagc actttgtaga aacagtttct ttggctggct cttacagaga
ttggtcctat
4741 tcaggacagc gtacagagtt aggtgtgaa tttcttaage gtggtgacaa
aattgtgtac
4801 cacactctgg agagccccgt cgagtttcat cttgacggtg aggttcttct
acttgacaaa
4861 ctaagagctc tcttatccct gcgggaggtt aagactataa aagtgttcac
aactgtggac
4921 aacactaatc tccacacaca gcttgtggat atgtctatga catatggaca
gcagtttggg
4981 ccaacatact tggatggtgc tgatgttaca aaaattaaac ctcatgtaaa
tcatgagggg
5041 aagactttct ttgtactacc tagtgatgac acactacgta gtgaagcttt
cgagtactac
5101 catactcttg atgagagttt tcttggtagg tacatgtctg ctttaaacca
cacaaagaaa
5161 tggaatttct ctcaagttgg tggtttaact tcaattaaat gggctgataa
caattgttat
5221 ttgtctagtg ttttattagc acttcaacag cttgaagtca aattcaatgc
accagcactt
5281 caagaggctt attatagagc ccgtgctggg gatgctgcta acttttgtgc
actcatactc
5341 gcttacagta ataaaactgt tggcgagctt ggtgatgtca gagaaactat
gacctatctt
5401 ctacagcatg ctaatttggg atctgcaaag cgagttctta atgtggtgtg
taaacattgt
5461 ggtcagaaaa ctactacctt aacgggtgta gaagctgtga tgtatatggg
tactctatct
5521 tatgataatc ttaagacagg tgtttccatt ccatgtgtgt gtggctgtga
tgctacacaa
5581 tatctagtac aacaagagtc ttcttttggg atgatgtctg caccacctgc
tgagtataaa
5641 ttacagcaag gtacattctt atgtgcgaat gagtacactg gtaactatca
gtgtgggcat
5701 tacactcata taactgctaa ggagaccctc tatcgtattg acggagctca
ccttcaaaag
5761 atgtcagagt acaaaggacc agtgactgat gttttctaca aggaaacatc
ttacactaca
5821 accatcaagc ctgtgtcgta taaactcgat ggagttactt acacagagat
tgaacaaaa
5881 ttggatgggt attataaaaa ggataatgct tactatacag agcagcctat
agaccttgta
5941 ccaactcaac cattacaaaa tgcgagtttt gataatttca aactcacatg
ttctaacaca
6001 aaatttctgt atgatttaaa tcaaatgaca ggcttcacaa agccagcttc
acgagagcta
6061 tctgtcacat tcttcccaga cttgaatggc gatgtagtgg ctattgacta
tagacactat

6121 tcagcgagtt tcaagaaagg tgctaaatta ctgcataagc caattgtttg
gcacattaac
6181 caggctacaa ccaagacaac gttcaaacca aacacttggg gtttacgttg
tctttggagt
6241 acaaagccag tagatacttc aaattcattt gaagttcttg cagtagaaga
cacacaagga
6301 atggacaatc ttgcttgtga aagtcaacaa cccacctctg aagaagtagt
ggaaaatcct
6361 accatacaga aggaagtcac agagtgtgac gtgaaaacta ccgaagttgt
aggcaatgtc
6421 atacttaaac catcagatga aggtgttaaa gtaacacaag agttaggtca
tgaggatctt
6481 atggctgctt atgtggaaaa cacaagcatt accattaaga aacctaata
gctttcacta
6541 gccttaggtt taaaaacaat tgccactcat ggtattgctg caattaatag
tgttccttgg
6601 agtaaaattt tggcttatgt caaaccattc ttaggacaag cagcaattac
aacatcaaat
6661 tgcgctaaga gattagcaca acgtgtgttt aacaattata tgccttatgt
gtttacatta
6721 ttgttccaat tgtgtacttt tactaaaagt accaattcta gaattagagc
ttcactacct
6781 acaactattg ctaaaaatag tgttaagagt gttgctaaat tatgtttgga
tgccggcatt
6841 aattatgtga agtcacccaa attttctaaa ttgttcacaa tcgctatgtg
gctattgttg
6901 ttaagtattt gcttaggttc tctaactctg gtaactgctg cttttggtgt
actcttatct
6961 aattttggtg ctccttctta ttgtaatggc gttagagaat tgtatcttaa
ttcgtctaac
7021 gttactacta tggatttctg tgaaggttct ttctcttgca gcatttgttt
aagtggatta
7081 gactcccttg attcttatcc agctcttgaa accattcagg tgacgatttc
atcgtaacag
7141 ctgacttga caatttttag tctggccgct gagtgggttt tggcatatat
gttgttcaca
7201 aaattctttt atttattagg tctttcagct ataatgcagg tgttctttgg
ctattttgct
7261 agtcatttca tcagcaattc ttggctcatg tggtttatca ttagtattgt
acaaatggca
7321 cccgtttctg caatggttag gatgtacatc ttctttgctt ctttctacta
catatggaag
7381 agctatgttc atatcatgga tggttgcacc tcttcgactt gcatgatgtg
ctataagcgc
7441 aatcgtgcca cgcggttga gtgtacaact attgttaatg gcatgaagag
atctttctat
7501 gtctatgcaa atggaggccg tggcttctgc aagactcaca attggaattg
tctcaattgt
7561 gacacatttt gcaactgtag tacattcatt agtgatgaag ttgctcgtga
tttgtcactc
7621 cagtttaaaa gaccaatcaa cctactgac cagtcatcgt atattgttga
tagtggtgct
7681 gtgaaaaatg ggcgcttca cctctacttt gacaaggctg gtcaaaagac
ctatgagaga
7741 catccgctct cccattttgt caatttagac aatttgagag ctaacaacac
taaaggttca
7801 ctgcctatta atgtcatagt ttttgatggc aagtccaaat gcgacgagtc
tgcttctaag
7861 tctgttcttg tgtactacag tcagctgatg tgccaacctt ttctgttgct
tgaccaagct
7921 cttgtatcag acgttgaga tagtactgaa gtttccgtta agatgtttga
tgcttatgtc
7981 gacacctttt cagcaacttt tagtgttcct atggaaaaac ttaaggcact
tgttgctaca
8041 gctcacagcg agttagcaaa ggggtgtagc ttagatgggtg tcctttctac
attcgtgtca
8101 gctgccgac aaggtgttgt tgataccgat gttgacacaa aggatgttat
tgaatgtctc
8161 aaactttcac atcactctga cttagaagtg acaggtgaca gttgtaacaa
tttcatgctc
8221 acctataata aggttgaaaa catgacgccc agagatcttg gcgcatgtat
tgactgtaat

8281 gcaaggcata tcaatgcccc agtagcaaaa agtcacaatg ttctactcat
ctggaatgta
8341 aaagactaca tgtctttatc tgaacagctg cgtaaacaaa ttcgtagtgc
tgccaagaag
8401 aacaacatac ctttttagact aacttggtgt acaactagac aggttggtcaa
tgtcataact
8461 actaaaatct cactcaaggg tggttaagatt gttagtactt gttttaaact
tatgcttaag
8521 gccacattat tgtgcgttct tgctgcattg gtttggtata tegtatatgcc
agtacatata
8581 ttgtcaatcc atgatgggta cacaatgaa atcattgggt acaaagccat
tcaggatggg
8641 gtcaactcgt acatcatttc tactgatgat tgttttgcaa ataaacatgc
tgggttttgac
8701 gcatgggtta gccagcgtgg tggttcatac aaaaatgaca aaagctgccc
tgtagtagct
8761 gctatcatta caagagagat tggtttcata gtgcctggct taccgggtac
tgtgctgaga
8821 gcaatcaatg gtgacttctt gcattttcta cctcgtggtt ttagtgctgt
tggcaacatt
8881 tgctacacac cttccaaact cattgagtat agtgattttg ctacctctgc
ttgcgttctt
8941 gctgctgagt gtacaatttt taaggatgct atgggcaaac ctgtgccata
ttggttatgac
9001 actaatttgc tagaggggtc tatttcttat agtgagcttc gtccagacac
tegttatgtg
9061 cttatggatg gttccatcat acagtttctt aacacttacc tggaggggtc
tgtagtagta
9121 gtaacaactt ttgatgctga gtactgtaga catggtacat gcgaaaggtc
agaagtaggt
9181 atttgcctat ctaccagtgg tagatgggtt ctttaataatg agcattacag
agctctatca
9241 ggagttttct gtggtgttga tgcgatgaat ctcatagcta acatctttac
tctcttctgt
9301 caacctgtgg gtgctttaga tgtgtctgct tcagtagtgg ctggtggtat
tattgccata
9361 ttggtgactt gtgctgccta ctactttatg aaattcagac gtgttttttg
tgagtacaac
9421 catgttgttg ctgctaatac acttttgttt ttgatgtctt tcactatact
ctgtctggtg
9481 ccagcttaca gctttctgcc gggagtctac tcagtctttt acttgtaact
gacattctat
9541 ttaccaatg atgtttcatt cttggctcac cttcaatggt ttgccatgtt
ttctctatt
9601 gtgccttttt ggataacagc aatctatgta ttctgtattt ctctgaagca
ctgccattgg
9661 ttctttaaca actatcttag gaaaagagtc atgtttaatg gagttacatt
tagtaccttc
9721 gaggaggctg ctttgtgtac ctttttgcgc aacaaggaaa tgaacctaaa
attgcgtagc
9781 gagacactgt tgccacttac acagtataac aggtatcttg ctctatataa
caagtacaag
9841 tatttcagtg gagccttaga tactaccagc tatcgtgaag cagcttgctg
ccacttagca
9901 aaggctctaa atgaacttag caactcaggt gctgatgttc tctaccaacc
accacagaca
9961 tcaatcactt ctgctgttct gcagagtggg tttaggaaaa tggcattccc
gtcaggcaaa
10021 gttgaagggt gcatggtaca agtaacctgt ggaactacaa ctcttaatgg
attgtggttg
10081 gatgacacag tatactgtcc aagacatgct atttgcacag cagaagacat
gcttaatcct
10141 aactatgaag atctgctcat tcgcaaatcc aacctagct ttcttgttca
ggctggcaat
10201 gttcaacttc gtgttattgg ccattctatg caaaattgtc tgettaggct
taaagttgat
10261 acttctaacc ctaagacacc caagtataaa tttgtccgta tccaacctgg
tcaaacattt
10321 tcagttctag catgctacaa tggttcacca tctgggtgtt atcagtgtgc
catgagacct
10381 aatcatacca ttaaagggtc tttccttaat ggatcatgtg gtagtggttg
ttttaacatt

10441 gattatgatt gcgtgtcttt ctgctatatg catcatatgg agcttccaac
aggagtacac
10501 gctgggtactg acttagaagg taaattctat ggtccatttg ttgacagaca
aactgcacag
10561 gctgcaggta cagacacaac cataacatta aatgttttgg catggctgta
tgctgctggt
10621 atcaatggtg atagggtggt tcttaataga ttcaccacta ctttgaatga
ctttaacctt
10681 gtggcaatga agtacaacta tgaacctttg acacaagatc atgttgacat
attgggacct
10741 ctttctgctc aaacaggaat tgccgtctta gatatgtgtg ctgctttgaa
agagctgctg
10801 cagaatggta tgaatggctg tactatcctt ggtagcacta ttttagaaga
tgagtttaca
10861 ccatttgatg ttgttagaca atgctctggt gttaccttcc aaggttaagtt
caagaaaatt
10921 gtaagggca ctcactcattg gatgctttta actttcttga catcactatt
gattcttgtt
10981 caaagtacac agtgggcact gtttttcttt gtttacgaga atgctttctt
gccatttact
11041 cttgggtatta tggcaattgc tgcagtgtct atgctgcttg ttaagcataa
gcacgcattc
11101 ttgtgcttgt ttctgttacc ttctcttgca acagttgctt actttaatat
ggctctacatg
11161 cctgctagct ggggtgatgc tatcatgaca tggcttgaat tggetgacac
tagcttgtct
11221 ggttataggc ttaaggattg tgttatgtat gcttcagctt tagttttgct
tattctcatg
11281 acagctcgca ctgtttatga tgatgctgct agacgtgttt ggacactgat
gaatgtcatt
11341 acacttgctt acaaaagtcta ctatggtaat gcttttagatc aagctatttc
catgtgggccc
11401 ttagttattt ctgtaacctc taactattct ggtgtcgta cgaactatcat
gttttttagct
11461 agagctatag tgtttgtgtg tggtgagtat taccattgt tatttattac
tggcaacacc
11521 ttacagtgtg tcatgcttgt ttattgtttc ttaggctatt gttgctgctg
ctactttggc
11581 cttttctgtt tactcaaccg ttacttcagg cttactcttg gtgtttatga
ctacttggc
11641 tctacacaag aatttaggta tatgaactcc caggggcttt tgccctctaa
gagtagtatt
11701 gatgctttca agcttaacat taagttgttg ggtattggag gtaaaccatg
tatcaagggt
11761 gctactgtac agtctaaaat gtctgacgta aagtgcacat ctgtgggtact
gctctcggtt
11821 cttcaacaac tttagagtaga gtcatcttct aaattgtggg cacaatgtgt
acaactccac
11881 aatgatattc ttcttgcaaa agacacaact gaagctttcg agaagatggt
ttctcttttg
11941 tctgttttgc tatccatgca ggggtgctgta gacattaata ggttggtcga
ggaaatgctc
12001 gataaccgtg ctactcttca ggctattgct tcagaattta gttctttacc
atcatatgcc
12061 gcttatgcca ctgccagga ggcctatgag caggctgtag ctaatggtga
ttctgaagtc
12121 gttctcaaaa agttaagaa atctttgaat gtggctaaaat ctgagtttga
ccgtgatgct
12181 gccatgcaac gcaagttgga aaagatggca gatcaggcta tgacccaaat
gtacaaacag
12241 gcaagatctg aggacaagag ggcaaaagta actagtgcta tgcaaacat
gctcttcact
12301 atgcttagga agcttgataa tgatgcactt aacaacatta tcaacaatgc
gcgtgatggt
12361 tgtgttccac tcaacatcat accattgact acagcagcca aactcatggt
tgttgccct
12421 gattatggta cctacaagaa cacttgatg ggtaacacct ttacatatgc
atctgcactc
12481 tgggaaatcc agcaagttgt tgatgcggat agcaagattg ttcaacttag
tgaaattaac
12541 atggacaatt caccaaaattt ggcttggcct cttattgtta cagctctaag
agccaactca

12601 gctgttaaac tacagaataa tgaactgagt ccagtagcac tacgacagat
gtcctgtgcg
12661 gctggtacca cacaacacgc ttgtactgat gacaatgcac ttgctacta
taacaattcg
12721 aaggagggtta ggtttgtgct ggcattacta tcagaccacc aagatctcaa
atgggctaga
12781 ttccctaaga gtgatggtac aggtacaatt tacacagaac tggaccacc
ttgtagggtt
12841 gttacagaca caccaaaagg gcctaaagtg aaatacttgt acttcatcaa
aggcttaaac
12901 aacctaaata gaggtatggt gctgggcagt ttagctgcta cagtacgtct
tcaggctgga
12961 aatgctacag aagtacctgc caattcaact gtgctttcct tctgtgcttt
tgtagtagac
13021 cctgctaaag catataagga ttacctagca agtgaggagc aaccaatcac
caactgtgtg
13081 aagatgttgt gtacacacac tggtagagga caggcaatta ctgtaacacc
agaagctaac
13141 atggaccaag agtcctttgg tggtagctca tgttgtctgt attgtagatg
ccacattgac
13201 catccaaatc ctaaggatt ctgtgacttg aaaggtaagt acgtccaaat
acctaccact
13261 tgtgctaattg acccagtggg ttttacctt agaaacacag tctgtaccgt
ctgcggaatg
13321 tggaaagggt atggctgtag ttgtgaccaa ctccggaac ccttgatgca
gtctgcggat
13381 gcatcaacgt ttttaaacgg gtttgcggtg taagtgcagc ccgtcttaca
ccgtgcggca
13441 caggcactag tactgatgtc gtctacaggg cttttgatat ttacaacgaa
aaagttgctg
13501 gttttgcaa gttcctaaaa actaattgct gtcgcttcca ggagaaggat
gagggaaggca
13561 atttattaga ctcttacttt gtagttaaga ggcatactat gtctaactac
caacatgaag
13621 agactattta taacttggtt aaagattgtc cagcgggtgc tgtccatgac
tttttcaagt
13681 ttagagtaga tggtagcatg gtaccacata tatcacgtca gcgtctaact
aaatacacia
13741 tggctgattt agtctatgct ctacgtcatt ttgatgaggg taattgtgat
acattaaaag
13801 aaatactcgt cacatacaat tgctgtgatg atgattattt caataagaag
gattggtatg
13861 acttcgtaga gaatcctgac atcttacgcg tatatgctaa cttagggtgag
cgtgtacgcc
13921 aatcattatt aaagactgta caattctgcg atgctatgcg tgatgcaggc
attgtaggcg
13981 tactgacatt agataatcag gatcttaatg ggaactggta cgatttcggt
gatttcgtac
14041 aagtagcacc aggtgcgga gttcctattg tggattcata ttactcattg
ctgatgcca
14101 tcctcacttt gactagggca ttggtgctg agtcccatat ggatgctgat
ctcgcaaaac
14161 cacttattaa gtgggatttg ctgaaatatg attttacgga agagagactt
tgtctcttcg
14221 accgttattt taaatattgg gaccagacat accatcccaa ttgtattaac
tgtttggatg
14281 atagggtgat ccttcattgt gcaaacctta atgtgttatt ttctactgtg
ttccaccta
14341 caagttttgg accactagta agaaaaatat ttgtagatgg tgttcctttt
gttgtttcaa
14401 ctggatacca ttttcgtgag ttaggagtcg tacataatca ggatgtaaac
ttacatagct
14461 cgcgtctcag tttcaaggaa cttttagtgt atgctgctga tccagctatg
catgcagctt
14521 ctggcaattt attgctagat aaacgcacta catgcttttc agtagctgca
ctaacaacaa
14581 atgttgcttt tcaaacgtgc aaaccggta attttaataa agacttttat
gactttgctg
14641 tgtctaaagg tttctttaag gaaggaaagt ctggtgaact aaaacacttc
ttctttgctc
14701 aggatggcaa cgctgctatc agtgattatg actattatcg ttataatctg
ccaacaatgt

14761 gtgatatcag acaactccta ttcgtagttg aagttgttga taaatacttt
gattgttacg
14821 atgggtggctg tattaatgcc aaccaagtaa tcgttaacaa tctggataaa
tcagctgggt
14881 tcccatttaa taaatggggg aaggctagac tttattatga ctcaatgagt
tatgaggatc
14941 aagatgcact tttcgcgtat actaagcgta atgtcatccc tactataact
caaatgaatc
15001 ttaagtatgc cattagtgc aagaatagag ctgcacccgt agctgggtgc
tctatctgta
15061 gtactatgac aaatagacag tttcatcaga aattattgaa gtcaatagcc
gccactagag
15121 gagctactgt ggtaattgga acaagcaagt tttacggtag ctggcataat
atgttaaaaa
15181 ctgtttacag tgatgtagaa actccacacc ttatgggttg ggattatcca
aaatgtgaca
15241 gagccatgcc taacatgctt aggataatgg cctctcttgt tcttgctcgc
aaacataaca
15301 cttgctgtaa cttatcacac cgtttctaca ggtagctaa cgagtgtgcg
caagtattaa
15361 gtgagatggg catgtgtggc ggctcactat atgttaaacc aggtggaaca
tcctccgggtg
15421 atgctacaac tgcttatgct aatagtgtct ttaacatttg tcaagctgtt
acagccaatg
15481 taaatgcact tctttcaact gatggtaata agatagctga caagtatgtc
cgcaatctac
15541 aacacaggct ctatgagtgt ctctatagaa atagggatgt tgatcatgaa
ttcgtgggatg
15601 agttttacgc ttacctgcgt aaacatttct ccatgatgat tctttctgat
gatgcggttg
15661 tgtgctataa cagtaactat gcggctcaag gtttagtagc tagcattaag
aactttaagg
15721 cagttcttta ttatcaaat aatgtgttca tgtctgaggc aaaatgttgg
actgagactg
15781 accttactaa aggacctcac gaattttgct cacagcatat aatgctagtt
aaacaaggag
15841 atgattacgt gtacctgctt taccagatc catcaagaat attaggcgca
ggctgttttg
15901 tcgatgatat tgcataaaca gatggtacac ttatgattga aagggtcgtg
tcactggcta
15961 ttgatgctta ccacttaca aaacatccta atcaggagta tgcgatgtc
tttcaacttg
16021 atttacaata cattagaaag ttacatgatg agcttactgg ccacatgttg
gacatgtatt
16081 ccgtaatgct aactaatgat aacacctcac ggtactggga acctgagttt
tatgaggcta
16141 tgtacacacc acatacagtc ttgcaggctg taggtgcttg tgtattgtgc
aattcacaga
16201 cttcacttcg ttgcgggtgc tgtattagga gaccattcct atgttgcaag
tgcgtctatg
16261 accatgtcat ttcaacatca cacaatttag tgttgctctg taatccctat
gtttgcaatg
16321 cccaggttg tgatgtcact gatgtgacac aactgtatct aggaggtatg
agctattatt
16381 gcaagtcaca taagctctcc attagttttc cattatgtgc taatggtcag
gtttttgggt
16441 tatacaaaaa cacatgtgta ggcagtgaca atgtcactga cttcaatgcg
atagcaacat
16501 gtgattggac taatgctggc gattacatac ttgccaacac ttgtactgag
agactcaagc
16561 ttttcgcagc agaaacgctc aaagccactg aggaaacatt taagctgtca
tatggatttg
16621 ccactgtacg cgaagtactc tctgacagag aattgcatct ttcattggag
gttggaatac
16681 ctagaccacc attgaacaga aactatgtct ttactgggta ccgtgtaact
aaaaatagta
16741 aagtacagat tggagagtac acctttgaaa aaggtgacta tgggtgatgt
gttggtgtaca
16801 gaggtactac gacatacaag ttgaatgttg gtgattactt tgtgttgaca
tctcacactg
16861 taatgccact tagtgcacct actctagtgc cacaagagca ctatgtgaga
attactggct

16921 tgtaccaaac actcaacatc tcagatgagt tttctagcaa tgttgcaaat
tatcaaaagg
16981 tcggcatgca aaagtactct acactccaag gaccacctgg tactggtaag
agtcattttg
17041 ccateggact tgctctctat taccatctg ctgcgatagt gtatacggca
tgctctcatg
17101 cagctggtga tgccttatgt gaaaaggcat taaaatattt gcccatagat
aatgtagta
17161 gaatcatacc tgcgcgtgcg cgcgtagagt gttttgataa attcaaagt
aattcaaac
17221 tagaacagta tgttttctgc actgtaaatg cattgccaga aacaactgct
gacattgtag
17281 tctttgatga aatctctatg gctactaatt atgacttgag tgttgtcaat
gctagacttc
17341 gtgcaaaaca ctacgtctat attggcgatc ctgctcaatt accagcccc
cgcacattgc
17401 tgactaaagg cacactagaa ccagaatatt ttaattcagt gtgcagactt
atgaaaacaa
17461 taggtccaga catgttcctt ggaactgtc gccgttgtec tgcgtgaaatt
gttgacactg
17521 tgagtgtctt agtttatgac aataagctaa aagcacacaa ggataagtca
gctcaatgct
17581 tcaaaatgtt ctacaaagggt gttattacac atgatgtttc atctgcaatc
aacagacctc
17641 aaataggcgt tgtaagagaa tttcttacac gcaatcctgc ttggagaaaa
gctgttttta
17701 tctcacctta taattcacag aacgctgtag cttcaaaaat cttaggattg
cctacgcaga
17761 ctgttgattc atcacagggt tctgaatatg actatgtcat attcacacaa
actactgaaa
17821 cagcacactc ttgtaatgtc aaccgcttca atgtggctat cacaagggca
aaaattggca
17881 ttttgtgcat aatgtctgat agagatcttt atgacaaact gcaatttaca
agtctagaaa
17941 taccacgtcg caatgtggct acattacaag cagaaaatgt aactggactt
tttaaggact
18001 gtagtaagat cttactggt cttcatccta cacaggcacc tacacacctc
agcgttgata
18061 taaagttcaa gactgaagga ttatgtgttg acataccagg cataccaaa
gacatgacct
18121 accgtagact catctctatg atgggtttca aaatgaatta ccaagtcaat
ggttacccta
18181 atatgtttat caccgcgaa gaagctattc gtcacgttcg tgcgtggatt
ggctttgatg
18241 tagagggtcg tcatgcaact agagatgctg tgggtactaa cctacctctc
cagctaggat
18301 tttctacagg tgttaactta gtagctgtac cgactggta tgttgacact
gaaaataaca
18361 cagaattcac cagagttaat gcaaaacctc caccagggtga ccagtttaaa
catcttatac
18421 cactcatgta taaaggcttg ccctggaatg tagtgcgat taagatagta
caaagtctca
18481 gtgatacact gaaaggattg tcagacagag tcgtgttcgt cctttggcg
catggctttg
18541 agcttacatc aatgaagtac tttgtcaaga ttggacctga aagaacgtgt
tgtctgtgtg
18601 acaaacgtgc aacttgcttt tctacttcat cagatactta tgctgtctgg
aatcattctg
18661 tgggttttga ctatgtctat aaccattta tgattgatgt tcagcagtgg
ggctttacgg
18721 gtaaccttca gagtaacct gaccaacatt gccaggta tggaaatgca
catgtggcta
18781 gttgtgatgc tatcatgact agatgtttag cagtccatga gtgctttgtt
aagcgcgttg
18841 attggctctgt tgaataccct attataggag atgaactgag ggttaattct
gcttgacaga
18901 aagtacaaca catggttgtg aagtctgcat tgcttgctga taagtttcca
gttcttcatg
18961 acattggaaa tccaaaggct atcaagtgtg tgcttcaggc tgaagtagaa
tggaagtctt
19021 acgatgctca gccatgtagt gacaaagctt acaaaataga ggaactcttc
tattcttatg

19081 ctacacatca cgataaatcc actgatgggtg tttgtttgtt ttggaattgt
aacgttgatc
19141 gttaccacagc caatgcaatt gtgtgtaggt ttgacacaag agtettgtca
aacttgaact
19201 taccaggctg tgatgggtgt agtttgtatg tgaataagca tgcattccac
actccagctt
19261 tcgataaaaag tgcatttact aatttaaagc aattgccttt cttttactat
tctgatagtc
19321 cttgtgagtc tcatggcaaa caagtagtgt eggatattga ttatgttcca
ctcaaactctg
19381 ctacgtgtat tacacgatgc aatttaggtg gtgctgtttg cagacaccat
gcaaattgagt
19441 accgacagta cttggatgca tataatatga tgatttctgc tggatttagc
ctatggattt
19501 acaaacaatt tgatacttat aacctgtgga atacatttac caggttacag
agttagaaa
19561 atgtggctta taatgttgtt aataaaggac actttgatgg acacgccggc
gaagcacctg
19621 tttccatcat taataatgct gtttacacaa aggtagatgg tattgatgtg
gagatctttg
19681 aaaataagac aacacttcct gttaatgttg catttgagct ttgggctaag
cgtaacatta
19741 aaccagtgcc agagattaag atactcaata atttgggtgt tgatatcgct
gctaatactg
19801 taatctggga ctacaaaaga gaagccccag cacatgtatc tacaataggt
gtctgcacaa
19861 tgactgacat tgccaagaaa cctactgaga gtgcttgttc ttcacttact
gtcttgtttg
19921 atggtagagt ggaaggacag gtagacctt ttagaaacgc ccgtaatggt
gttttaataa
19981 cagaagggtc agtcaaagg ctaacacctt caaagggacc agcacaagct
agcgtcaatg
20041 gagtcacatt aattggagaa tcagtaaaaa cacagttaa ctactttaag
aaagtagacg
20101 gcattattca acagttgctt gaaacctact ttactcagag cagagactta
gaggatttta
20161 agcccagatc acaaatggaa actgactttc tcgagctgc tatggatgaa
ttcatacagc
20221 gatataagct cgagggttat gccttcgaac acatcgttta tggagatttc
agtcatggac
20281 aacttggcgy tcttcattta atgataggct tagccaagcg ctcacaagat
tcaccactta
20341 aattagagga ttttatccct atggacagca cagtgaaaaa ttacttcata
acagatgcgc
20401 aaacagggtc atcaaaatgt gtgtgttctg tgattgatct tttacttgat
gactttgtcg
20461 agataataaa gtcacaagat ttgtcagtga tttcaaaagt ggtcaagggt
acaattgact
20521 atgctgaaat ttcattcatg ctttgggtga aggatggaca tgttgaaacc
ttctacccaa
20581 aactacaagc aagtcaagcg tggcaaccag gtgttgcgat gcctaacttg
tacaagatgc
20641 aaagaatgct tcttgaaaag tgtgaccttc agaattatgg tgaaaatgct
gttatacca
20701 aaggaataat gatgaatgtc gcaaagtata ctcaactgtg tcaatactta
aatacactta
20761 ctttagctgt accctacaac atgagagtta ttcactttgg tctggctct
gataaaggag
20821 ttgcaccagg tacagctgtg ctcagacaat ggttgccaac tggcacacta
cttgtcgatt
20881 cagatcttaa tgacttcgtc tccgacgcag attctacttt aattggagac
tgtgcaacag
20941 tacatacggc taataaatgg gaccttatta ttagcgatat gtatgacct
aggaccaaac
21001 atgtgacaaa agagaatgac tctaaagaag ggtttttcac ttatctgtgt
ggatttataa
21061 agcaaaaact agccctgggt ggttctatag ctgtaaagat aacagagcat
tcttggaaatg
21121 ctgaccttta caagcttat ggccatttct catgggtggac agcttttgtt
acaaatgtaa
21181 atgcatcatc atcggaagca tttttaattg gggctaacta tcttggcaag
ccgaaggaa

21241 aaattgatgg ctataccatg catgctaact acattttctg gaggaacaca
aatcctatcc
21301 agttgtcttc ctattcactc ttgacatga gcaaatttcc tcttaaatta
agaggaactg
21361 ctgtaatgtc tcttaaggag aatcaaatca atgatatgat ttattctctt
ctggaaaaag
21421 gtaggcttat cattagagaa aacaacagag ttgtggtttc aagtgatatt
cttgtaaca
21481 actaaacgaa catgtttatt ttcttattat ttcttactct cactagtgg
agtgaacctg
21541 accgggtgcac cacttttgat gatgttcaag ctctaatta cactcaacat
acttcactca
21601 tgaggggggt ttactatcct gatgaaattt ttagatcaga cactctttat
ttaactcagg
21661 atttatttct tccattttat tctaattgta cagggtttca tactattaat
catacgtttg
21721 gcaaccctgt catacctttt aaggatggta tttattttgc tgccacagag
aatcaaatg
21781 ttgtccgtgg ttgggttttt gggtctacca tgaacaacaa gtcacagtcg
gtgattatta
21841 ttaacaattc tactaatgtt gttatacag catgtaactt tgaattgtgt
gacaaccctt
21901 tctttgctgt ttctaaacct atgggtacac agacacatac tatgatattc
gataatgcat
21961 ttaattgcac ttctgagtac atatctgat ccttttcgct tgatgtttca
gaaaagtcag
22021 gtaattttta acacttacga gagtttgtgt ttaaaaataa agatgggttt
ctctatgttt
22081 ataagggtca tcaacctata gatgtagttc gtgatctacc ttctggtttt
aacactttga
22141 aaactatttt taagtgtcct ctgtgtatta acattacaaa ttttagagcc
attcttacag
22201 ccttttcacc tgctcaagac atttggggca cgtcagctgc agcctatttt
gttggtctatt
22261 taaagccaac tacatttatg ctcaagtatg atgaaaatgg tacaatcaca
gatgctgttg
22321 attgtttctca aaatccactt gctgaactca aatgctctgt taagagcttt
gagattgaca
22381 aaggaattta ccagacctct aatttcaggg ttgtccctc aggagatgtt
gtgagattcc
22441 ctaatattac aaacttgtgt ccttttggag aggtttttta tgctactaaa
ttcccttctg
22501 tctatgcatg ggagagaaaa aaaatttcta attgtgttgc tgattactct
gtgctctaca
22561 actcaacatt tttttcaacc tttaagtgtc atggcggttc tgccactaag
ttgaatgac
22621 ttgtcttctc caatgtctat gcagattctt ttgtagtcaa gggagatgat
gtaagacaaa
22681 tagcgccagg acaaactggt gttattgtc attataatta taaattgcca
gatgatttca
22741 tgggttgtgt ccttgcttgg aatactagga acattgatgc tacttcaact
ggtaattata
22801 attataaata taggtatctt agacatggca agcttaggcc ctttgagaga
gacatatcta
22861 atgtgccttt cccccctgat ggcaaacctt gcacccacc tgctcttaat
tgttattggc
22921 cattaaatga ttatggtttt tacaccacta ctggcattgg ctaccaacct
tacagagttg
22981 tagtactttc ttttgaactt ttaaatgcac cggccacggg ttgtggacca
aaattatcca
23041 ctgaccttat taagaaccag tgtgtcaatt ttaattttta tggactcact
ggactgggtg
23101 tgtaactcc ttcttcaaag agatttcaac catttcaaca atttggcgt
gatgtttctg
23161 atttcaactga ttccgttcga gatcctaaaa catctgaaat attagacatt
tcaccttgcg
23221 cttttggggg tgtaagtgt attacacctg gaacaaatgc ttcactgtga
gttgctgttc
23281 tatatcaaga tgtaactgc actgatgttt ctacagcaat tcatgcagat
caactcacac
23341 cagcttggcg catatattct actggaacaa atgtattcca gactcaagca
ggctgtctta

23401 taggagctga gcattgtcgac atttcttatg agtgcgacat tcctattgga
gctggcattt
23461 gtgctagtta ccatacagtt tctttattac gtagtactag ccaaaaatct
attgtgggctt
23521 atactatgtc tttaggtgct gatagtcaa ttgcttactc taataacacc
attgctatac
23581 ctactaactt ttcaattagc attactacag aagtaatgcc tgtttctatg
gctaaaaacct
23641 ccgtagattg taatatgtac atctgcggag attctactga atgtgcta
ttgcttctcc
23701 aatatggtag cttttgcaca caactaaatc gtgcactctc aggtattgct
gctgaacagg
23761 atcgcaacac acgtgaagtg ttgctcaag tcaaacaaat gtacaaaacc
ccaactttga
23821 aatatttttg tggttttaat ttttcacaaa tattacctga ccctctaaag
ccaactaaga
23881 ggtcttttat tgaggacttg ctctttaata aggtgacact cgtgatgct
ggcttcatga
23941 agcaatatgg cgaatgccta ggtgatatta atgctagaga tctcatttgt
gcgcaagaat
24001 tcaatggact tacagtgttg ccacctctgc tcaactgatga tatgattgct
gctacactg
24061 ctgctctagt tagtggtagt gccactgctg gatggacatt tggtagctggc
gctgctcttc
24121 aaataccttt tgctatgcaa atggcatata ggttcaatgg cattggagtt
acccaaaatg
24181 ttctctatga gaacaaaaa caaatcgcca accaatttaa caaggcgatt
agtcaaatc
24241 aagaatcact tacaacaaca tcaactgcat tgggcaagct gcaagacgtt
gttaaccaga
24301 atgctcaagc attaaacaca cttgttaaac aacttagctc taattttggt
gcaatttcaa
24361 gtgtgctaaa tgatctctt tcgcgacttg ataaagtcga ggcggaggta
caaattgaca
24421 ggtaattac aggcagactt caaagccttc aaacctatgt aacacaacaa
ctaactcagg
24481 ctgctgaaat cagggtctt gctaactctg ctgctactaa aatgtctgag
tgtgttcttg
24541 gacaatcaaa aagagttgac ttttgggaa agggctacca ccttatgtcc
ttcccacaag
24601 cagccccgca tgggtgtgtc ttctacatg tcacgtatgt gccatcccag
gagaggaact
24661 tcaccacagc gccagcaatt tgtcatgaag gcaaagcata ctccctcgt
gaaggtgtt
24721 ttgtgtttaa tggcacttct tggtttatta cacagaggaa cttcttttct
ccacaaataa
24781 ttactacaga caatacatt gtctcaggaa attgtgatgt cgttattggc
atcataaaca
24841 acacagttta tgatctctg caacctgagc ttgactcatt caaagaagag
ctggacaagt
24901 acttcaaaaa tcatacatca ccagatgttg atcttggcga catttcaggc
attaacgctt
24961 ctgtcgtcaa cattcaaaaa gaaattgacc gcctcaatga ggtcgtctaa
aattttaaag
25021 aatcactcat tgaccttcaa gaattgggaa aatatgagca atatattaaa
tggccttgg
25081 atgtttggct cggcttcatt gctggactaa ttgccatcgt catggttaca
atcttgcttt
25141 gttgcatgac tagttgttgc agttgcctca aggggtgcatg ctcttggtg
tcttgctgca
25201 agtttgatga ggatgactct gagccagttc tcaagggtgt caaattacat
tacacataaa
25261 cgaacttatg gatttgttta tgagattttt tactcttaga tcaattactg
cacagccagt
25321 aaaaattgac aatgcttctc ctgcaagtac tgttcatgct acagcaacga
taccgttaca
25381 agcctcactc ccttcggat ggcttggtat tggcgttgca tttcttgctg
tttttcagag
25441 cgctacaaa ataattgcgc tcaataaaaag atggcagcta gccctttata
agggcttcca
25501 gttcatttgc aatttactgc tgctatttgt taccatctat tcacatctt
tgcttgctgc

25561 tgcaggtatg gaggcgcaat ttttgtacct ctatgccttg atatattttc
tacaatgcat
25621 caacgcatgt agaattatta tgagatgttg gctttgttgg aagtgc aaat
ccaagaacccc
25681 attactttat gatgccaaact actttgtttg ctggcacaca cataactatg
actactgtat
25741 accatataac agtgtcacag atacaattgt cgttactgaa ggtgacggca
tttcaacacc
25801 aaaactcaaa gaagactacc aaattggtgg ttattctgag gataggcact
caggtgttaa
25861 agactatgtc gttgtacatg gctatttcac cgaagtttac taccagcttg
agtctacaca
25921 aattactaca gacactggta ttgaaaatgc tacattcttc atctttaaca
agcttgttaa
25981 agaccaccg aatgtgcaaa tacacacaat cgacggctct tcaggagtgtg
ctaattccagc
26041 aatggatcca atttatgatg agccgacgac gactactagc gtgcctttgt
aagcacaaga
26101 aagtgagtag gaacttatgt actcattcgt ttcggaagaa acaggtagct
taatagttaa
26161 tagcgtactt ctttttcttg ctttcgtggg attcttgcta gtcacactag
ccatccttac
26221 tgcgcttcga ttgtgtgcgt actgctgcaa tattgttaac gtgagtttag
taaaaccaac
26281 ggtttacgtc tactcgctg ttaaaaatct gaactcttct gaaggagtgc
ctgatcttct
26341 ggtctaaacg aactaactat tattattatt ctgtttgga ctttaacatt
gcttatcatg
26401 gcagacaacg gtactattac cgttgaggag cttaacaac tcttgaaca
atggaacctt
26461 gtaataggtt tcctattcct agcctggatt atgttactac aatttgctta
ttctaactcg
26521 aacaggtttt tgtacataat aaagcttgtt ttctctggc tcttgtggcc
agtaacactt
26581 gcttgttttg tgcttgcgtc tgtctacaga attaatggg tgactggcgg
gattgcgatt
26641 gcaatggctt gtattgtagg cttgatgtgg cttagctact tcgttgcttc
cttcaggctg
26701 ttgtctgta ccgcgtcaat gtggctatc aaccagaaa caaacattct
tctcaatgtg
26761 cctctccggg ggacaattgt gaccagaccg ctcatggaaa gtgaacttgt
cattggtgct
26821 gtgatcttc gtggctactt gcgaatggcc ggacactccc tagggcgctg
tgacattaag
26881 gacctgccaa aagagatcac tgtggctaca tcacgaacgc tttcttatta
caaattagga
26941 gcgtcgcagc gtgtaggcac tgattcaggt ttgctgcat acaaccgcta
ccgtattgga
27001 aactataaat taaatacaga ccacgccggt agcaacgaca atattgcttt
gctagtacag
27061 taagtgacaa cagatgtttc atcttgttga cttccagggtt acaatagcag
agatattgat
27121 tatcattatg aggactttca ggattgctat ttggaatctt gacgttataa
taagttcaat
27181 agtgagacaa ttatttaagc ctctaactaa gaagaattat tcggagtttag
atgatgaaga
27241 acctatggag tttagattac cataaaacga acatgaaaat tattctcttc
ctgacattga
27301 ttgtatttac atcttgcgag ctatatcact atcaggagtg tgttagaggt
acgactgtac
27361 tactaaaaga accttgccca tcaggaacat acgagggcaa ttcaccattt
caccctcttg
27421 ctgacaataa atttgacta acttgacta gcacacactt tgcttttgct
tgtgctgacg
27481 gtactcgaca tacctatcag ctgcgtgcaa gatcagtttc accaaaactt
ttcatcagac
27541 aagaggaggt tcaacaagag ctctactcgc cactttttct cattgttgct
gctctagtat
27601 ttttaatact ttgcttcacc attagagaa agacagaatg aatgagctca
ctttaattga
27661 cttctatttg tgctttttag ctttctgct attccttggt ttaataatgc
ttatttatatt

27721 ttgggttttca ctcgaaatcc aggatctaga agaaccttgt accaaagtct
aaacgaacat
27781 gaaactttctc attgttttga cttgtatttc tctatgcagt tgcatatgca
ctgtagtaca
27841 gcgctgtgca tctaataaac ctcatgtgct tgaagatcct tgtaagggtac
aacactaggg
27901 gtaatactta tagcactgct tggctttgtg ctctaggaaa ggttttacct
tttcatagat
27961 ggcacactat ggttcaaaca tgcacaccta atgttactat caactgtcaa
gatccagctg
28021 gtggtgcgct tatagctagg tgttgggtacc ttcataaagg tcaccaaact
gctgcattta
28081 gagacgtact tgttgtttta aataaacgaa caaattaaaa tgtctgataa
tggaccoccaa
28141 tcaaaccaac gtagtgcgcc ccgcattaca tttggtggac ccacagattc
aactgacaat
28201 aaccagaatg gaggacgcaa tggggcaagg ccaaacagc gccgacccca
agggtttacc
28261 aataatactg cgtcttggtt cacagctctc actcagcatg gcaaggagga
acttagattc
28321 cctcagaggc agggcggttc aatcaacacc aatagtgttc cagatgacca
aattggctac
28381 taccgaagag ctaccgcagc agttcgtggt ggtgacggca aaatgaaaga
gctcagcccc
28441 agatggtact tctattacct aggaactggc ccagaagctt cacttcccta
cggcgctaac
28501 aaagaaggca tcgtatgggt tgcaactgag ggagccttga atacacccaa
agaccacatt
28561 ggcaccgcga atcctaataa caatgctgcc accgtgctac aacttcctca
aggaacaaca
28621 ttgccccaaag gctctacgc agaggaagc agaggcgga gtcaagcctc
ttctcgctcc
28681 tcatacgtg gtgcgggtta ttcaagaaat tcaactcctg gcagcagtag
gggaaattct
28741 cctgctcgaa tggctagcgg aggtggtgaa actgccctcg cgctattgct
gctagacaga
28801 ttgaaccagc ttgagagcaa agtttctggt aaaggccaac aacaacaagg
ccaaactgtc
28861 actaagaaat ctgctgctga ggcactctaa aagcctcgcc aaaaacgtac
tgccacaaaa
28921 cagtacaacy tcaactcaagc atttgggaga cgtgggtccag aacaaacca
aggaaatttc
28981 ggggaccaag acctaatcag acaaggaaact gattacaaac attggccgca
aattgcacaa
29041 tttgctccaa gtgcctctgc attctttgga atgtcacgca ttggcatgga
agtcacacct
29101 tcgggaacat ggctgactta tcattggagc attaaattgg atgacaaaga
tcacaaattc
29161 aaagacaacy tcatactgct gaacaagcac attgacgcat acaaacatt
cccaccaaca
29221 gagcctaaaa aggacaaaaa gaaaaagact gatgaagctc agcctttgcc
gcagagacaa
29281 aagaagcagc ccactgtgac tcttcttctc gcggctgaca tggatgattt
ctccagacaa
29341 cttcaaaatt ccattgagtg agcttctgct gattcaactc aggcataaac
actcatgatg
29401 accacacaag gcagatgggc tatgtaaacg ttttcgcaat tccgtttacg
atacatagtc
29461 tactcttgtg cagaatgaat tctcgtaact aaacagcaca agtaggttta
gttaacttta
29521 atctcacata gcaatcttta atcaatgtgt aacattaggg aggacttgaa
agagccacca
29581 cattttcctc gaggccacgc ggagtacgat cgagggtaca gtgaataatg
ctaggagagag
29641 ctgcctatat ggaagagccc taatgtgtaa aattaatttt agtagtgcta
tccccatgtg
29701 attttaatag cttcttagga gaatgacaaa aaaaaaaaaa aaaaaaaaaa a
//

SARS-associated Coronavirus

More Information on SARS

- US Centre for Disease Control
- The World Health Organisation

About

About

News

Employment

Faculty

Contacts

Activities

Cancer Genomics

Affymetrix Platform

Functional Genomics

Gene Regulation

Gene Expression

Physical Mapping

Sequencing

Custom Services

Bioinformatics

Genome BC/Genome

Canada

Software

UBIC

Training

Bioinformatics

Bioinformatics

Workshops

VanBLG

Resources

AccPrimer

C. elegans

ciRED

Ensembl

ExperimentBank.org

GeneCards

Human Map

Mouse Atlas of Gene

Expression

ProteBLAST

SRS Server

Human BAC library

Search

Go

The Genome Sciences Centre SARS associated virus genome sequencing project is collaboration with the British Columbia Centre for Disease Control and the National Microbiology Laboratory, Winnipeg.

Clone Availability

Clones derived from the SARS associated virus are now available from the [SARS clone page](#).

Genomic Sequence Availability

2003/04/29 | The finished whole genome assembly is available for [download](#). The fasta file contains the finished genome totalling 29,751 base pairs.

2003/04/14 | The SARS-associated Coronavirus (TOR2) sequence has been deposited in the NCBI Genbank database under accession [AY274119](#).

2003/04/12 | The current genome shotgun data is available for [download](#). The format is raw sequence reads in fasta format, and not vector or quality clipped.

2003/04/12 | The draft whole genome assembly is available for [download](#). The fasta file contains the first draft genome assembly totalling 29,736 base pairs. This assembly may contain errors.

News Stories



The Linux Journal features a cover article about our lab's use of Linux in the sequencing of the SARS virus.

The cover of the magazine shows a 3-D image of a BLASTX comparison of the genome of the SARS Tor2 Coronavirus against the known proteins of other similar viruses. The image was generated using the Linux version of the Sockeye 3-D genome viewer, developed in our lab.

A feature story, "B.C. lab cracks suspected SARS code", is currently featured on [cbs.ca](#)

VANCOUVER - Scientists at a Canadian research centre say they've mapped the

Download Sequence

The shotgun sequence reads of the TOR isolate are available for download.

DOWNLOAD SEQUENCE

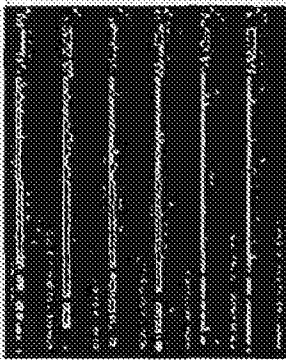
The fasta file contains 3,080 sequence reads totalling 3,621,777 base pairs.

The finished genome sequence assembly release 3, of the TOR2 isolate is available for download.

DOWNLOAD SEQUENCE

The fasta file contains the first draft genome assembly totalling 29,751 base pairs. The assembly may contain errors.

Comparison to other Nidovirales



This image (zoom) was generated by the Sockeye viewer and shows black areas the protein complement of the nidoviral family against build 7.4 of the SARS-

Collaborator Access



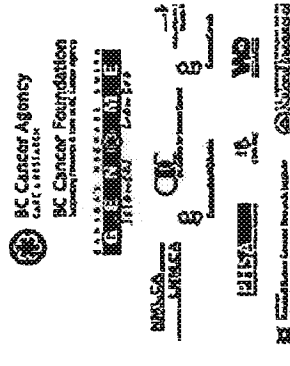
Text Size: A A A A

genomic sequence of the virus believed to cause SARS, a move that could take the world a step closer to a reliable test for the illness [more](#)

More information about SARS can be accessed [here](#).

There is significant dissimilarity in the 5' of the polyprotein region where the leaf protein of several other Coronaviruses is located. There is also significant variation in the structural proteins at the 3' end.

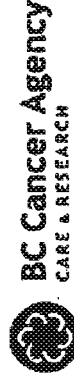
Participants



Logos of participating organizations.

BC Cancer Agency news releases

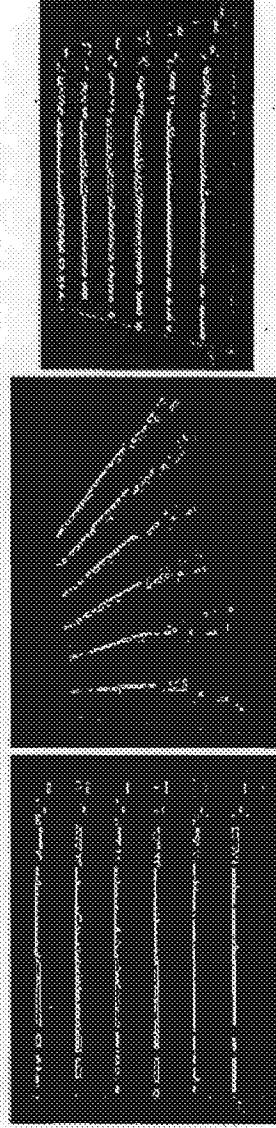
- 2003/04/12 | Genome Sciences Centre sequences SARS associated corona virus [more](#)
- 2003/04/07 | Genome Sciences Centre in partnership to sequence the SARS virus genome [more](#)



BC Cancer Agency
CARE & RESEARCH

Comparison to other Nidovirales

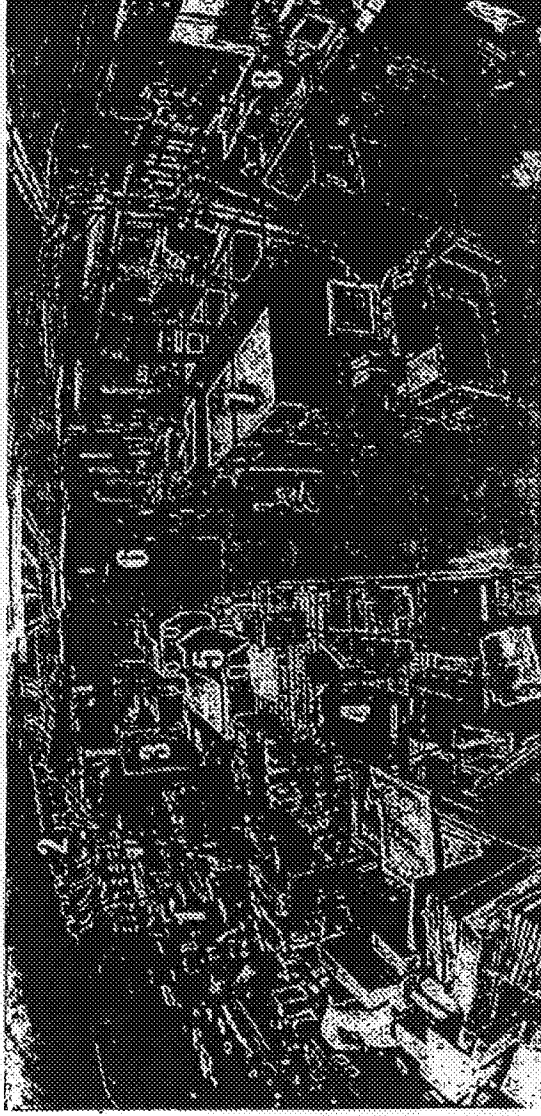
These images were generated by the Sockeye genome viewer, developed at the Genome Sciences Centre. The images show the comparison of the sequence of the TOR2 isolate to the sequence of other related nidovirales. The tracks show blastx comparison to (1) avian infectious bronchitis virus, (2) bovine coronavirus, (3) human coronavirus 229E, (4) murine hepatitis virus, (5) porcine epidemic diarrhoea virus, (6) transmissible gastroenteritis virus. An annotated image is also available.



Images courtesy of the GSC bioinformatics team.

Around the Lab

The Genome Sciences Centre carries a wide range of basic and applied research in genomics, including sequencing, physical mapping, functional genomics, cancer genomics, gene expression and technology development. The photo below shows the sequencing area of the lab. Other images are available [here](#).



1 96- and 384-well liquid handling robotics automate preparation of sequencing reactions.

2 The lab's power and modularity requirements are unique. Network and power connections drop down from the ceiling to facilitate arrangement of instruments.

3 DNA and other samples are stored in large freezers which maintain a constant temperature of -80 degrees Celsius.

4 Thermocyclers are used to activate sequencing reactions by exposing the samples to a specific cycle of heating and cooling.

5 Robotic colony pickers are used to isolate and collect bacterial colonies which are used to amplify the DNA to be sequenced.

6 Linux-based Beowulf cluster used for data analysis and sequence assembly builds.

7 State-of-the-art robotic sequencers include the ABI 3700, ABI 3730XL, Megabace 1000 and Megabace 4000. Sequencers are loaded with multiple 96- or 384-well plates and are capable of producing over 300,000 quality base pairs from a 384-well plate per run.

8 96- and 384-head liquid handling robots used to dispense reaction liquids to multiple plates.



2003/04/14 | Images from today's media conference



Images of people and equipment around the lab.

Contacts Email: webmaster@bcgsc.ca Phone: 604 877-6084

Powered by Plone

gb AAV97992.1.1	spike glycoprotein [SARS coronavirus B033]	2562	0.0	
gb AAU04662.1.1	spike glycoprotein [SARS coronavirus civet019]	2561	0.0	
gb AAV97999.1.1	spike glycoprotein [SARS coronavirus B012]	2560	0.0	
gb AAV98900.1.1	spike glycoprotein [SARS coronavirus C025]	2560	0.0	
gb AAV97984.1.1	spike glycoprotein [SARS coronavirus A001]	2560	0.0	
gb AAU04649.1.1	spike glycoprotein [SARS coronavirus civet010]...	2560	0.0	
gb AAV98901.1.1	spike glycoprotein [SARS coronavirus C028]	2559	0.0	
gb AAV97998.1.1	spike glycoprotein [SARS coronavirus C018]	2558	0.0	
gb AAU04646.1.1	spike glycoprotein [SARS coronavirus civet007]	2557	0.0	
gb AAV91621.1.1	spike glycoprotein [SARS coronavirus A022] >gb...	2556	0.0	
gb ACJ60703.1.1	spike glycoprotein [recombinant coronavirus]	2223	0.0	
ref YP_001364399.1.1	spike protein [Bat SARS CoV Rm1/2004] >gb...	2068	0.0	
sp Q8Q478.1 SPIN_E0279	RecName: Full=Spike glycoprotein; Sho...	2061	0.0	
ref YP_001362361.1.1	spike protein [Bat SARS CoV Rp3/2004] >sp...	2059	0.0	G
ref YP_001362373.1.1	spike glycoprotein [Bat coronavirus HKU3]...	2049	0.0	G
gb AA241029.1.1	spike glycoprotein [bat SARS coronavirus HKU3-2]	2047	0.0	
gb AA241029.1.1	spike protein [Bat coronavirus (BtCoV/273/2005)]	2009	0.0	
ref YP_001362385.1.1	spike protein [Bat SARS CoV Rf1/2004] >gb...	2006	0.0	
gb AAU93320.1.1	spike glycoprotein [SARS coronavirus GZ0403]	1558	0.0	
gb AAE77236.1.1	S1 protein [SARS coronavirus GD322]	1460	0.0	
gb AAU25930.1.1	S2 protein [SARS coronavirus GD322]	1103	0.0	
gb AAW56618.1.1	truncated SARS coronavirus spike glycoprotein ...	1013	0.0	
gb AAS44761.1.1	spike glycoprotein [SARS coronavirus TW-HP1] >...	927	0.0	
gb AAS44782.1.1	spike glycoprotein [SARS coronavirus TW-YM2]	926	0.0	
gb AAS44779.1.1	spike glycoprotein [SARS coronavirus TW-GD3] >...	921	0.0	
gb AAS44781.1.1	spike glycoprotein [SARS coronavirus TW-GD4]	920	0.0	
gb AAS91092.1.1	spike glycoprotein [SARS coronavirus CUHK-L2]	765	0.0	
gb ABN10919.1.1	Spike glycoprotein [Bat coronavirus HKU9-2]	639	0.0	
ref YP_001029971.1.1	spike glycoprotein [Bat coronavirus HKU9]...	632	1e-178	G
gb ABN10927.1.1	Spike glycoprotein [Bat coronavirus HKU9-3]	622	6e-176	
gb ABN10935.1.1	Spike glycoprotein [Bat coronavirus HKU9-4]	606	6e-171	
ref YP_001029962.1.1	spike glycoprotein [Bat coronavirus HKU5]...	591	2e-166	G
ref YP_729204.1.1	spike protein [Bat coronavirus (BtCoV/133/20...]	583	3e-164	G
gb AAS48579.1.1	spike glycoprotein [SARS coronavirus cw049]	580	3e-163	
gb ABN10984.1.1	Spike glycoprotein [Bat coronavirus HKU5-2]	579	7e-163	
gb ABN10993.1.1	Spike glycoprotein [Bat coronavirus HKU5-3] >g...	578	2e-162	
gb AEG11962.1.1	spike [Bat coronavirus (BtCoV/A434/2005)]	575	2e-161	
gb ABN10952.1.1	Spike glycoprotein [Bat coronavirus HKU4-3]	572	9e-161	
gb ABN10948.1.1	Spike glycoprotein [Bat coronavirus HKU4-2]	572	1e-160	
ref YP_001029993.1.1	spike glycoprotein [Bat coronavirus HKU4]...	572	1e-160	G

Alignments Select All Get selected sequences Distance tree of results Multiple alignment

>ref|NP_828851.1| **G** E2 glycoprotein precursor [SARS coronavirus]
 gb|AAP41037.1| spike glycoprotein [SARS coronavirus Tor2]
 Length=1255

GENE ID: 1489668 S | E2 glycoprotein precursor [SARS coronavirus]
 (Over 10 PubMed links)

Score = 2602 bits (6745), Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 1255/1255 (100%), Positives = 1255/1255 (100%), Gaps = 0/1255 (0%)

Query	1	MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFL	60
Sbjct	1	MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFL	60
Query	61	PFYSNVTGFHTINHTFGNPVIFPKDGIYFAATEKSNVVRGWVFGSTMNNKSQS VIIINNS	120
Sbjct	61	PFYSNVTGFHTINHTFGNPVIFPKDGIYFAATEKSNVVRGWVFGSTMNNKSQS VIIINNS	120
Query	121	TNVVIRACNFELCDNPFPAVSKPMGTQHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFK	180
Sbjct	121	TNVVIRACNFELCDNPFPAVSKPMGTQHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFK	180
Query	181	HLREFVFKNKDGLYVYKGYQPLDVVRDLPSGFNTLKPFIKPLGINITNFRAILTAFSP	240

10/822,303
 Attachment (3)
 SEQIDNO:6042 vs
 Tor2

Sbjct	181	HLREFVFKNKDGLYVYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFSP	240
Query	241	AQDIWGTSAAYFVGYLKPTTFMLKYDENGTTITDAVDCSQNPLAELKCSVKSFEDKGIY	300
Sbjct	241	AQDIWGTSAAYFVGYLKPTTFMLKYDENGTTITDAVDCSQNPLAELKCSVKSFEDKGIY	300
Query	301	QTSNFRVVPBGDVVRFPNITNLCPPFGEVFNATKFPSSVYAWERKKISNCVADYSVLNSTF	360
Sbjct	301	QTSNFRVVPBGDVVRFPNITNLCPPFGEVFNATKFPSSVYAWERKKISNCVADYSVLNSTF	360
Query	361	FSTFKCYGVSATKLNLCFSNVYADSFVVGDDVRQIAPGQTGVIADYNYKLPDDFMGCV	420
Sbjct	361	FSTFKCYGVSATKLNLCFSNVYADSFVVGDDVRQIAPGQTGVIADYNYKLPDDFMGCV	420
Query	421	LAWNTRNIDATSTGNYNKYRYLRHGKLRPFERDISNVFSPDGKPCPTPALNCYWPLND	480
Sbjct	421	LAWNTRNIDATSTGNYNKYRYLRHGKLRPFERDISNVFSPDGKPCPTPALNCYWPLND	480
Query	481	YGFYTTTGIGYQPYRVVLSFELLNAPATVCGPKLSTDLIKNCVNFNFNGLTGTGVLTTP	540
Sbjct	481	YGFYTTTGIGYQPYRVVLSFELLNAPATVCGPKLSTDLIKNCVNFNFNGLTGTGVLTTP	540
Query	541	SSKRFQPFQFGRDVSDFTDVSRDPKTSSEILDISPCAFGGVSVITPGTNASSEVAVLYQD	600
Sbjct	541	SSKRFQPFQFGRDVSDFTDVSRDPKTSSEILDISPCAFGGVSVITPGTNASSEVAVLYQD	600
Query	601	VNCTDVSTAIHADQLTPAWRIYSTGNVFPQTQAGCLIGAETHVDTSEYCDIPIGAGICASY	660
Sbjct	601	VNCTDVSTAIHADQLTPAWRIYSTGNVFPQTQAGCLIGAETHVDTSEYCDIPIGAGICASY	660
Query	661	HTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMKTSVDC	720
Sbjct	661	HTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMKTSVDC	720
Query	721	NMYICGDSTECANLLQYGSFCTQLNRLSGIAAEQDRNTREVFAQVKQMYKTPTLKYFG	780
Sbjct	721	NMYICGDSTECANLLQYGSFCTQLNRLSGIAAEQDRNTREVFAQVKQMYKTPTLKYFG	780
Query	781	GFNFSQILPDPLKPTKRSFIEDLLFNKVTADAGFMKQYGECLGDINARDLICAQKFNGL	840
Sbjct	781	GFNFSQILPDPLKPTKRSFIEDLLFNKVTADAGFMKQYGECLGDINARDLICAQKFNGL	840
Query	841	TVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTONVLYE	900
Sbjct	841	TVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTONVLYE	900
Query	901	NQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLN	960
Sbjct	901	NQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLN	960
Query	961	DILSRDLKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSK	1020
Sbjct	961	DILSRDLKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSK	1020
Query	1021	RVDFCGKGHYLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVVFVN	1080
Sbjct	1021	RVDFCGKGHYLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVVFVN	1080
Query	1081	GTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEELDKEYFN	1140
Sbjct	1081	GTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEELDKEYFN	1140
Query	1141	HTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWL	1200
Sbjct	1141	HTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWL	1200
Query	1201	GFIAGLIAIVMTILLCCMTSCCSCLKGACSCGSCCKFDEDDSEPVLKGVKLHYT	1255
Sbjct	1201	GFIAGLIAIVMTILLCCMTSCCSCLKGACSCGSCCKFDEDDSEPVLKGVKLHYT	1255

>gb|AAR07630.1| spike glycoprotein [SARS coronavirus BJ302]
Length=1255

Score = 2600 bits (6739), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 1253/1255 (99%), Positives = 1254/1255 (99%), Gaps = 0/1255 (0%)

Query	1	MFIFLLFLTLTSGSGLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLTYLTQDLFL	60
Sbjct	1	MFIFLLFLTLTSGSGLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLTYLTQDLFL	60
Query	61	PFYSNVTGFHTINHTFGNPVPIPKDGIYFAATEKSNVVRGWVFGSTMNNKSQSIIINNS	120


[CDC Home](#)
[Search](#)
[Health Topics A-Z](#)

Office of Enterprise Communication

Phone (404)639-3286

**MEDIA
RELATIONS**
[Media Home](#) | [Contact Us](#)

Press Release

Attachment (4)
10/822,303

For Immediate Release

April 14, 2003
 Contact: CDC Media Relations:
 404-639-3286

Media Relations Links

- [About Us](#)
- [Contact Us](#)
- [Frequently Asked Questions](#)
- [Media Site Map](#)

CDC News

- [Press Release Library](#)
- [Transcripts](#)
- [MMWR Summaries](#)
- [B-Roll Footage](#)
- [Upcoming Events](#)

Related Links

- [Centers at CDC](#)
- [Data and Statistics](#)
- [Health Topics A-Z](#)
- [Image Library](#)
- [Publications, Software and Other Products](#)
- [Global Health Odyssey](#)

[Find your state or local health department](#)
[HHS News](#)
[National Health Observances](#)


Div. of Media Relations
 1600 Clifton Road
 MS D-14
 Atlanta, GA 30333
 (404) 639-3286
 Fax (404) 639-7394

CDC Lab Sequences Genome of New Coronavirus

The Centers for Disease Control and Prevention (CDC) announced today that it has sequenced the genome for the coronavirus believed to be responsible for the global epidemic of severe acute respiratory syndrome or SARS. The CDC sequence is nearly identical to that determined by a Canadian laboratory late last week. The significant difference is that the CDC-determined sequence has 15 additional nucleotides, which provides the important beginning of the sequence, CDC scientists said.

The results came just 12 days after a team of 10 scientists, supported by numerous technicians, began working around the clock to grow cells taken from a throat culture taken from one of the SARS patients in Vero cells (African green monkey kidney cells) in order to reproduce the ribonucleic acid (RNA) of the disease-causing coronavirus. The new sequence has 29,727 nucleotides, which places it well within the typical RNA boundaries for coronaviruses. Members of this viral family tend to have between 29,000 and 31,000 nucleotides.

Identifying the genetic sequence of a new virus is important to efforts to treat or prevent it, said Dr. Julie Gerberding, CDC director. "Research laboratories can use this information to begin to target antiviral drugs, to form the basis for developing vaccines, and to develop diagnostic tests that can lead to early detection."

In sequencing the genome, CDC scientists worked closely with coronavirus experts at academic institutions across the United States. "This is an active, working community of scientific experts who have been contributing their knowledge and expertise throughout this investigation," said William Bellini, Ph.D., SARS laboratory team coordinator.

The nearly identical findings in the US and Canada are important because they were derived from different individuals who were infected in different countries. This suggests that the virus probably originated from a common source.

The CDC's analysis of the virus is far from finished, officials emphasized. Because coronaviruses have the ability to mutate rapidly, scientists will compare the sequences from viruses isolated in cell culture to those obtained from diseased tissues taken from SARS patients. "This is essentially a draft. Now we need to see if what we have identified in the laboratory matches what's causing disease in patients," Bellini said.

But the groundbreaking work of isolating the genomic sequence speeds

the task of comparison.

###

CDC protects people's health and safety by preventing and controlling diseases and injuries; enhances health decisions by providing credible information on critical health issues; and promotes healthy living through strong partnerships with local, national, and international organizations.

[Media Home Page](#) | [Accessibility](#) | [Privacy Policy](#) | [Contact Us](#)

[CDC Home](#) | [Search](#) | [Health Topics A-Z](#)

This page last updated April 14, 2003
URL: <http://www.cdc.gov/od/oc/media/pressrel/r030414.htm>

[United States Department of Health and Human Services](#)
[Centers for Disease Control and Prevention](#)
Office of Communication
Division of Media Relations